HYDROLOGICAL PROCESSES
Hydrol. Process. **17**, 3387–3409 (2003)
Published online 21 August 2003 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hyp.1299

An evaluation of sediment rating curves for estimating suspended sediment concentrations for subsequent flux calculations

Arthur J. Horowitz*
U.S. Geological Survey, Suite 130, Peachtree Business Center, 3039 Amwiler Road, Atlanta, GA 30360, USA

Abstract:

In the absence of actual suspended sediment concentration (SSC) measurements, hydrologists have used sediment rating (sediment transport) curves to estimate (predict) SSCs for subsequent flux calculations. Various evaluations of the sediment rating-curve method were made using data from long-term, daily sediment-measuring sites within large (>1 000 000 km^2), medium (<1 000 000 to >1000 km^2), and small (<1000 km^2) river basins in the USA and Europe relative to the estimation of suspended sediment fluxes. The evaluations address such issues as the accuracy of flux estimations for various levels of temporal resolution as well as the impact of sampling frequency on the magnitude of flux estimation errors.

The sediment rating-curve method tends to underpredict high, and overpredict low SSCs. As such, the range of errors associated with concomitant flux estimates for relatively short time-frames (e.g. daily, weekly) are likely to be substantially larger than those associated with longer time-frames (e.g. quarterly, annually) because the over- and underpredictions do not have sufficient time to balance each other. Hence, when error limits must be kept under ±20%, temporal resolution probably should be limited to quarterly or greater.

The evaluations indicate that over periods of 20 or more years, errors of <1% can be achieved using a single sediment rating curve based on data spanning the entire period. However, somewhat better estimates for the entire period, and markedly better annual estimates within the period, can be obtained if individual annual sediment rating curves are used instead. Relatively accurate (errors <±20%) annual suspended sediment fluxes can be obtained from hydrologically based monthly measurements/samples. For 5-year periods or longer, similar results can be obtained from measurements/samples collected once every 2 months. In either case, hydrologically based sampling, as opposed to calendar-based sampling is likely to limit the magnitude of flux estimation errors. Published in 2003 by John Wiley & Sons, Ltd.

KEY WORDS sediment rating curves; suspended sediment concentration; suspended sediment flux estimates

INTRODUCTION

Since the 1970s, there has been growing interest in estimating the fluvial transport of suspended sediment. The reasons are numerous and diverse, and include such issues as: contaminant transport, water-quality trends, reservoir sedimentation, channel and harbour silting, soil erosion and loss, as well as ecological and recreational impacts (Walling, 1977; Ferguson, 1986; de Vries and Klavers, 1994; Horowitz, 1995; Horowitz *et al*., 2001). A further impetus, at least in the USA, stems from the need to determine total maximum daily loads (TMDLs) for sediment, as well as for many sediment-associated constituents, under the requirements of the Clean Water Act of 1972. Nearly 17% of all currently required TMDLs deal with excess sediment or its presumptive biological impact (Keyes and Radcliffe, 2002).

The calculation of fluxes or loads requires both discharge and concentration data (e.g. de Vries and Klavers, 1994; Phillips *et al*., 1999). Typically continuous, or near-continuous, discharge data can be calculated from

* Correspondence to: Arthur J. Horowitz, U.S. Geological Survey, Suite 130, Peachtree Business Center, 3039 Amwiler Road, Atlanta, GA 30360, USA. E-mail: horowitz@usgs.gov

Received 25 November 2002
Accepted 2 January 2003

in situ devices such as a stage recorder. Stage is then converted to discharge based on a site-specific stage–discharge relationship. On the other hand, the vast majority of suspended sediment concentration (SSC) data typically result from manually collected individual samples taken at fixed temporal intervals; occasionally, the fixed interval samples are supplemented by event samples. More recently, as a result of the presumed linkage between SSC and discharge, SSC data have been generated on the basis of hydrologically based sampling rather than on fixed interval (calendar based) sampling (e.g. Horowitz, 1995; Horowitz *et al*., 2001). Although both approaches can provide a representation of the site-specific range of SSC, the hydrological approach will do so more rapidly than the fixed interval approach; however, samples are still collected manually.

During the past 20 years, continuous or near-continuous SSC data have been generated by using automatic samplers, or by measuring applicable surrogates such as turbidity (e.g. Walling, 1977; Horowitz, 1995). Whereas the requisite equipment (e.g. *in situ* turbidometers (turbidity probes), automatic samplers) for these newer sampling/measurement procedures is relatively inexpensive to obtain, installation and equipment calibration, as well as operational and maintenance costs are fairly high. Further, site-specific calibrations are needed to generate cross-sectionally representative data (Horowitz, 1995; Christensen, 2001). Meaningful calibrations typically require from 20 to 30 sample sets (simultaneously collected depth- and width-integrated isokinetic samples paired with point samples/measurements) collected over a typical annual range of hydrological conditions in order to be effective (Horowitz, 1995; Christensen, 2001). Hence, currently continuous or near-continuous SSC data sets are still relatively rare.

For more than 60 years, in the absence of actual continuous or near-continuous SSC data, hydrologists have used rating (sediment transport) curves to estimate (predict) daily SSCs for flux calculations. Although there are more than 20 methods for developing sediment rating curves, the most common is a power function (regression) that relates SSC to water discharge, with the discharge measurement constituting the independent variable (e.g. de Vries and Klavers, 1994; Phillips *et al*., 1999; Asselman, 2000). This requires the log-transformation of SSC and discharge data prior to the analysis. Comparisons of actual and predicted SSC, partially as a result of scatter about the regression line, as well as the conversion of results from log-space to arithmetic-space, indicate that sediment rating curves can substantially underpredict actual concentrations (Ferguson, 1986; Walling and Webb, 1988; Asselman, 2000). To compensate, various method modifications have been applied. These modifications have included dividing the SSC/discharge data into seasonal or hydrological groupings, developing various correction factors, or using non-linear regression equations (Duan, 1983, Ferguson, 1986; Walling and Webb, 1988; de Vries and Klavers, 1994; Phillips *et al*., 1999; Asselman, 2000; Holtschlag, 2001).

In 1995–1996, the U.S. Geological Survey's (USGS) National Stream Quality Accounting Network (NASQAN), was revised from an occurrence-and-distribution-based network to a large-river flux-based water-quality monitoring network (Hooper *et al*., 2001; Horowitz *et al*., 2001). Suspended sediment concentrations were required to calculate fluxes for sediment, as well as for various sediment-associated constituents (e.g. trace elements, nutrients). Initially, plans called for the installation of turbidometers at each site to provide continuous or near-continuous data on SSC. These devices were to be calibrated during the first 3 years of the programme, when hydrologically based sampling would be most intense (Horowitz *et al*., 2001). Unfortunately, owing to various resource constraints, the turbidometers were never installed; hence, the requisite SSCs/fluxes had to be determined from site-specific sediment rating curves. Since 1996, the effect of using the sediment rating-curve approach, relative to such issues as sampling frequency, temporal resolution and errors associated with flux estimates continue to be evaluated (Horowitz *et al*., 2001).

Additional evaluations resulted from the author's involvement with a technical advisory group (TAG) established to develop guidelines for generating scientifically credible sediment TMDLs in Georgia (Keyes and Radcliffe, 2002). Lastly, a series of evaluations were generated to determine the most appropriate way to calculate annual suspended sediment and sediment-associated chemical fluxes covering the first 5 years of the NASQAN programme. Succinctly, the major technical issues are:

1. how accurately can fluxes be estimated for various levels of temporal resolution (e.g. daily, weekly, monthly, yearly) using sediment rating curves, or, stated another way, what are the error limits associated with sediment rating-curve generated flux estimates for various levels of temporal resolution;
2. what levels of sampling frequency are required to achieve certain levels of predictive accuracy?

On the other hand, the underlying management issues are how much will it cost, and how much error is tolerable while still permitting sound management decisions relative to sediment and sediment-related problems? Many of these technical comparisons/evaluations are presented and discussed herein.

METHODS

Within the NASQAN network, the Mississippi River at Thebes site is unique because it constitutes the only long-term, ongoing, daily sediment-measuring site in the network. As such, the data from this site are uniquely suited to evaluating such issues as sampling frequency, temporal resolution and estimation errors, in the context of rating curve-generated flux estimates. The initial sediment rating curve evaluations were performed using daily data from this site (Table I). As the evaluation proceeded, additional data sets were used (e.g. Rhine at Maxau, Green River at Munford, Table I). In the majority of cases, these data sets had to contain daily SSC and discharge measurements for periods of at least 20 years. More limited data sets were used to confirm some of the results from the larger sets (e.g. Mississippi River at Clinton, Missouri at Hermann, Table I).

Data generation

Although there are numerous formulae/methods for calculating suspended sediment fluxes, and there is disagreement over which formula provides the most accurate flux estimates, the requisite data components always are discharge and SSC (Porterfield, 1977; de Vries and Klavers, 1994; Horowitz, 1995, Phillips, *et al.*, 1999). All the SSCs and discharge measurements for the USA sites were collected by the USGS. Discharge was calculated based on a 'continuous' (typically, measurements are taken every 15–60 min) record of stage from site-specific stage–discharge relationships. However, for purposes of these calculations, mean daily discharge determined from the 'continuous' data was used. For the long-term ($\geq$20 years) daily sediment stations, site-specific techniques used for determining SSC were as follows.

1. <u>Mississippi River at Thebes</u>: the SSC from a single depth-integrated vertical sample collected near the centroid of flow was related by a regression equation to a manually collected cross-sectional isokinetic depth- and width-integrated sample with subsequent laboratory determinations of SSC.
2. <u>Yadkin River at Yadkin College</u>: the SSC from a single depth-integrated vertical sample was used; the location of the vertical was adjusted annually such that the laboratory determined SSC from this non-representative sample was equivalent to the SSC from a representative isokinetic depth- and width-integrated sample collected from the entire cross-section.
3. <u>Green River at Munford</u>: see Yadkin River.
4. <u>Schuylkill River</u>: see Yadkin River.
5. <u>Rhine River at Maxau</u>: daily discharge and SSC were determined by the Bundesanstalt für Gewasserkunde (BfG); SSC was determined on samples collected 1 m below the surface (Deutscher Verband für Wasserwirtschaft und Kulturbau, 1986).

The SSC concentrations for all the additional (non-daily) sites used in these evaluations were determined from manually collected isokinetic depth- and width-integrated samples following standard USGS collection techniques (Edwards and Glysson, 1988). Discharge at these sites was determined in the same way as at the long-term daily site at Thebes.

Table I. Basic hydrological and sample data for the sites used in this study

Site	Basin area (km^2)	Dates	Number of samples	Discharge ($m^3 s^{-1}$)				Suspended sediment concentration ($mg l^{-1}$)			
				Minimum	Maximum	Mean	Median	Minimum	Maximum	Mean	Median
Mississippi River at Thebes, Illinois, USA	1 847 188	1/10/80–30/9/00	6 962	1 464	27 697	6 967	5 692	13	3 890	378	267
Rhine River at Maxau, Germany	50 200	31/10/73–30/10/93	7 305	462	4 210	1 274	1 189	1	236	25	21
Yadkin River at Yadkin College, North Carolina, USA	5 905	1/11/51–30/9/89	13 880	9·3	1 869	85·0	62·3	1	2 970	149	70
Green River at Munford, Kentucky, USA	4 333	1/10/69–30/9/89	6 985	4·4	1 778	85·0	45·3	1	2 720	84	38
Schuykill River at Berne, Pennsylvania, USA	919	1/10/61–30/9/81	7 305	1·5	736	19·8	12·5	1	558	9	5
Mississippi River at Clinton, Iowa, USA	221 704	10/12/74–30/9/00	191	402	6 712	1 671	1 359	2	281	65	51
Missouri River at Hermann, Missouri, USA	1 357 160	24/1/73–30/9/00	287	510	9 997	2 974	2 549	22	4 690	696	378
Ohio River near Grand Chain, Illinois, USA	526 029	14/2/73–30/9/00	287	864	34 267	8 468	5 919	3	485	85	58
Mississippi River at St Francisville, Louisiana, USA	2 913 750	5/6/78–30/9/00	241	3 144	41 914	14 811	12 914	30	711	216	198
Broad River near Bell, Georgia, USA	3 704	14/1/58–29/10/79	375	9·3	1 020	59·5	36·8	9	871	176	137
Oconee River at Dublin, Georgia, USA	11 396	2/3/61–13/12/71	312	12·6	1 900	144	87·8	7	911	61	51

Flux calculations

Daily sediment fluxes for the USA sites used in this study were calculated using the following formula

$$\text{suspended sediment flux (t day}^{-1}) = [Q(\text{ft}^3\ \text{s}^{-1})][\text{SSC}(\text{mg L}^{-1})][0{\cdot}00245] \quad (1)$$

where Q is discharge and SSC the suspended sediment concentration. Note that the unit of discharge is non-metric because that is the USA standard, but fluxes were converted to metric units (tonnes) through the use of an appropriate constant (0·00245) (Porterfield, 1977). On the other hand, daily sediment fluxes for the Rhine River site were calculated using the following formula

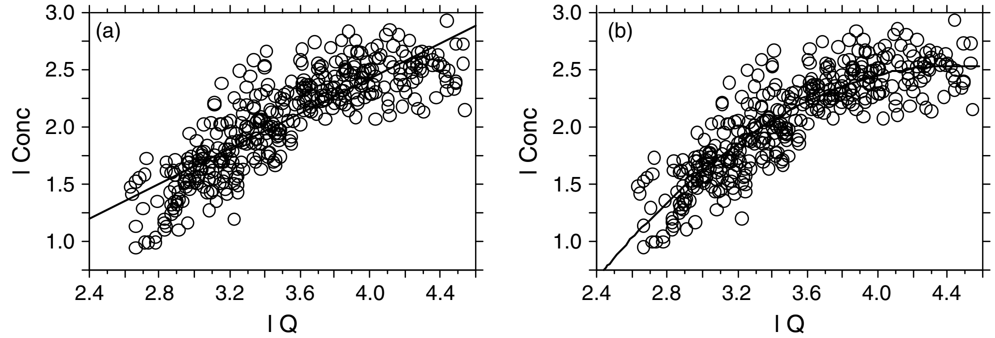
$$\text{suspended sediment flux (t day}^{-1}) = [Q(\text{m}^3\ \text{s}^{-1})][\text{SSC}(\text{mg L}^{-1})][0{\cdot}0864] \quad (2)$$

where Q is discharge and SSC the suspended sediment concentration. In this case, metric units were used throughout (Porterfield, 1977). Fluxes for periods longer than 1 day were calculated by summing the appropriate number of daily fluxes.

Sediment rating curves and flux estimations

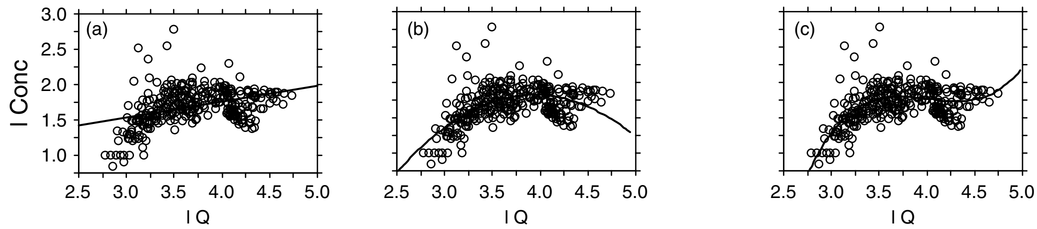
As noted previously, there are numerous methods for estimating SSC from discharge in the absence of actual SSC data (e.g. Walling and Webb, 1981; de Vries and Klavers, 1994; Phillips *et al*., 1999). These methods include extrapolation and interpolation, with potential additional modifications using various correction factors (Duan, 1983; Ferguson, 1986; Holtschlag, 2001). Further improvements have been noted when calibration data sets have been subdivided into seasonal (e.g. wet/dry) or hydrological (e.g. rising limb/falling limb) groupings (Walling, 1977; Hansen and Bray, 1993; Sichingabula, 1998; Asselman, 2000). The basis for the most commonly used procedure (extrapolation) is the determination of a log–log regression relating SSC to discharge. The efficacy of this approach is dependent on the sampling frequency. In other words, estimation accuracy depends on the number of paired data points available to develop the rating curve, and how well they represent the ranges of discharge and SSC at a site (Walling and Webb, 1981; Roberts, 1997). Despite the existence of these various methods, the NASQAN experience in estimating (predicting) SSCs for the calculation of annual or longer term suspended sediment fluxes indicated that typically, very good results (errors $\leq\pm15\%$) could be obtained using either linear or second-order polynomial sediment rating curves, with and without a 'smearing' correction (Horowitz *et al*., 2001).

The procedure used for the NASQAN programme to estimate daily mean SSCs is the one used for this series of comparisons/evaluations. Briefly, all the available (current and historic) discharge and SSC data for a specific site were used. The data are combined into a single calibration set and log transformed. A series of regression equations (e.g. linear, polynomial) are calculated and subsequently evaluated. Residual analyses are performed for each regression. If the pattern for the residual analysis is random, then predicted daily SSCs are generated for each of the acceptable curves. In turn, these are used to calculate daily fluxes for each set of data points from each curve. As can occur with these types of calculations, the conversion from log-space to arithmetic-space can produce a bias; when compensation is necessary to reduce/eliminate bias, a 'smearing correction' is applied (Duan, 1983). The daily fluxes, based on a daily mean SSC from each curve and a measured daily mean discharge, are then summed to produce a measure of total flux for a selected temporal resolution (e.g. weekly, monthly, quarterly). This approach can produce as many as four flux estimates: (i) linear uncorrected; (ii) linear 'smearing' corrected; (iii) polynomial uncorrected; and (iv) polynomial 'smearing' corrected. Each of the four results are compared with the total flux for the calibration set. The linear or polynomial rating curve, with or without a 'smearing' correction, which produces a total flux estimate closest to that for the calibration set is then used to predict SSCs where no actual sample data are available. An example for the Broad River near Bell, Georgia site indicates that SSC predictions



Y = -.651 + .77X	Y = -6.271 + 3.966X - .447X^2
n = 376	n = 376
r^2 = .70	r^2 = .76
Actual Flux = 1.56 Mt	Actual Flux = 1.56 Mt
Predicted Flux = 1.83 Mt	Predicted Flux = 1.45 Mt
Corrected Predicted Flux = 2.12 Mt	Corrected Predicted Flux = 1.64 Mt
Maximum Underprediction = -480 mg l^{-1}	Maximum Underprediction = -481 mg l^{-1}
Maximum Overprediction = 554 mg l^{-1}	Maximum Overprediction = 245 mg l^{-1}

Figure 1. Sediment rating curves for Broad River near Bell, Georgia, USA. (a) Linear regression and (b) second-order polynomial regression, and associated data. The labels on the graphs stand for log of discharge (l Q) and log of concentration (l Conc)



Y = .846 + .228X	Y = -6.358 + 4.155X - .528X^2	Y = -38.883 + 30.799X - 7.733X^2 + .644X^3
n = 311	n = 311	n = 311
r^2 = .15	r^2 = .32	r^2 = .38
Actual Flux = 0.38 Mt	Actual Flux = 0.38 Mt	Actual Flux = 0.38 Mt
Predicted Flux = 0.38 Mt	Predicted Flux = 0.34 Mt	Predicted Flux = 0.35 Mt
Corrected Predicted Flux = 0.45 Mt	Corrected Predicted Flux = 0.40 Mt	Corrected Predicted Flux = 0.40 Mt
Maximum Underprediction = -550 mg l^{-1}	Maximum Underprediction = -532 mg l^{-1}	Maximum Underprediction = -526 mg l^{-1}
Maximum Overprediction = 44 mg l^{-1}	Maximum Overprediction = 45 mg l^{-1}	Maximum Overprediction = 45 mg l^{-1}
Absolute Mean of Errors = 21 mg l^{-1}	Absolute Mean of Errors = 23 mg l^{-1}	Absolute Mean of Errors = 22 mg l^{-1}

Figure 2. Sediment rating curves for Oconee River at Dublin, Georgia, USA. (a) Linear regression, (b) second-order polynomial regression and (c) third-order polynomial regression, and associated data. The labels on the graphs stand for log of discharge (l Q) and log of concentration (l Conc)

from a corrected second-order polynomial rating curve would be used to estimate a daily mean SSC and to calculate suspended sediment fluxes (Figure 1).

In some cases, as for the Oconee River at Dublin, Georgia, the selection of an appropriate sediment rating curve can be more complex (Figure 2). If the sole deciding factor is the accuracy of the flux predictions, then an uncorrected linear regression provides the most accurate result (Figure 2a). The second- (Figure 2b) and third-order (Figure 2c) uncorrected polynomial regressions provide slightly lower estimates (8–11%), whereas the corrected polynomial regressions provide a slightly higher (+5%) estimate of flux. In fact, in this particular instance, an equivalent result to that obtained with the corrected polynomial regressions can be obtained simply by using the average value for SSC (59 mg L^{-1}; 0·40 Mt). The maximum over- and underpredictions for individual SSCs, for all three curves are about the same; albeit, the third-order polynomial curve has the smallest underprediction (Figure 2). In turn, the means of the absolute SSC prediction errors for all three curves also are about the same (*c*. 22%). On the other hand, the third-order polynomial curve provides the most accurate predictions for the lowest SSCs (Figure 2c). Hence, the selection of an appropriate sediment rating curve ultimately may rest on programmatic priorities.

The terms 'error', 'per cent error', or 'per cent difference', are used interchangeably in the text and tables to refer to the differences between measured SSCs and actual fluxes (calculated from measured SSCs and discharge), and predicted (estimated) SSCs derived from sediment rating curves and estimated fluxes (calculated from estimated SSCs and measured discharge). It should be understood that these terms do not encompass any of the potential errors associated with sampling, nor the subsequent laboratory determinations of SSC. Further, these terms do not include the potential errors associated with converting non-representative sample-derived SSCs to cross-sectional values through the use of either so-called 'box' equations nor adjusted sample locations (see site-specific techniques described above). The differences between measured and predicted values may be expressed as a concentration (mg L^{-1}) as in the case of SSCs, as a mass (tonnes) in the case of fluxes, or as a percentage for either measurement. In the latter case, percentage was calculated as follows, and is based on the view that the measured value is the 'standard'

$$\%\,\text{difference} = [(\text{predicted value}) - (\text{measured value})/(\text{measured value})] \times 100 \qquad (3)$$

A minus sign indicates an underprediction, whereas a positive sign indicates an overprediction relative to the actual (measured) value. As an example, consider a measured SSC of 1600 mg L^{-1} and a predicted value of 400 mg L^{-1}. Based on Equation (3), the difference would be −75%. On the other hand, if the measured SSC was 400 mg L^{-1} and the predicted value was 1600 mg L^{-1}, Equation (3) would indicate a difference of +300%.

All the linear and non-linear sediment rating curves, the SSC estimates for the unsampled periods, the calculation of an appropriate 'smearing' correction and all the associated statistical calculations (e.g. residual analyses) were performed on a Macintosh desktop computer using a commercially available statistics package (StatView 5·0; the use of brand names is for identification purposes only and does not constitute an endorsement by the U.S. Government). Subsequent daily and summary flux estimates were calculated on the same computer using an Excel spreadsheet program from Microsoft Office 98.

RESULTS AND DISCUSSION

When fluxes were estimated for the first three water years (1 October 1996 to 30 September 1998) of the revised NASQAN programme, sediment rating curves were generated from regression analyses based on site-specific calibration sets (discharge and SSC) that spanned the entire database. This included measurements dating back to the early 1970s because the sampling and analytical procedures for both the original and revised NASQAN programmes were considered identical (Horowitz *et al*., 2001). In contrast, when a second set of estimates was calculated for a CD-ROM publication covering the first five water years of the programme, they were generated using a predictive model called LOADEST (Crawford, 1996; Aulenbach B.T., USGS, personal communication, 2002). However, unlike the original regression equations, the model was calibrated using only data that were generated under the revised programme [1996–2000 water years (WY)]. Disagreements over the potential lack of comparability between the first and the second set of flux estimates, and the need for potential data caveats were the initial rationale for this comprehensive sediment rating curve evaluation.

The effect of calibration set composition on sediment rating curves and flux estimates

The actual suspended sediment flux for the Mississippi River at Thebes site, based on daily samples covering the first 5-year period of the revised NASQAN programme (1996–2000 WY) was 414 Mt (Table II). On the other hand, the estimated flux from a site-specific rating curve based on the same daily values was 409 Mt, a 1% underestimate (Table II). Despite this close agreement for the entire 5-year period, maximum errors in daily estimates of SSC ranged from −60% (predicted = 655 mg L^{-1}, actual = 1630 mg L^{-1}) to +146% (predicted = 861 mg L^{-1}, actual = 350 mg L^{-1}) (Table II).

Changes in the composition of the calibration sets used to develop the sediment rating curves for the Thebes site, and/or for calibrating the LOADEST model, produced significantly different estimates of cumulative suspended sediment fluxes covering the same 5-year period (1996–2000 WY) relative to the actual flux of 414 Mt (Table II). Limited calibration sets composed of different data (≤72 samples) generated during

Table II. The impact on 5-year suspended sediment flux estimates (water years 1996–2000) for the Mississippi River at Thebes NASQAN site as a result of variations in the composition of the calibration sets used to generate site-specific rating curves

Calculation	Calibration source	Flux (Mt)	Maximum (mg L^{-1})	Underestimate (%)	Maximum (mg L^{-1})	Overestimate (%)	Difference (Mt)	Percentage difference	r^2
Actual flux	All data[a]	414							
Predicted flux	All data[b]	409	−975	−60	511	146	−5	−1	0·73
Predicted flux	NASQAN data (F)[c]	439	−885	−54	645	184	25	6	0·74
Predicted flux	NASQAN data (F)-LOADEST[d]	450					36	9	
Predicted flux	NASQAN data (D)[e]	423	−896	−55	673	192	12	3	0·66
Predicted flux	Thebes data/NASQAN dates[f]	447	−932	−57	545	156	33	8	0·73
Predicted flux	96 WY data[g]	409	−964	−59	513	147	−5	−1	0·79
Predicted flux	97 WY data[h]	369	−1054	−65	395	160	−45	−11	0·64
Predicted flux	98 WY data[i]	381	−1060	−65	380	154	−33	−8	0·64
Predicted flux	99 WY data[j]	461	−862	−53	660	189	47	11	0·68
Predicted flux	00 WY data[k]	423	−969	−59	491	199	9	2	0·76
Predicted flux	15 year calibration[l]	512	−853	−52	611	247	98	24	0·53
Predicted flux	20 year calibration[m]	485	−889	−55	573	232	71	17	0·56

[a] Flux calculated from the daily mean values for SSC and discharge, $n = 1718$.
[b] Flux estimated from SSCs based on a regression-derived rating curve calibrated using daily mean values for SSC and discharge, $n = 1718$.
[c] Flux estimated from actual sample SSCs (determined by filtration, $n = 72$) and daily mean values for discharge using regression models.
[d] Flux estimated from actual sample SSCs (determined by filtration, $n = 72$) and daily mean values for discharge using LOADEST (Aulenbach B.T., USGS, personal communication, 2002).
[e] Flux estimated from actual sample SSCs (determined by dewatering, $n = 58$) and daily mean values for discharge using regression models.
[f] Flux estimated from daily sample SSCs (determined by filtration, $n = 72$) and daily mean values for discharge collected on the same dates as the NASQAN samples using regression models.
[g] Flux estimated from SSCs based on a regression-derived rating curve calibrated using daily mean values for SSC and discharge collected during the 1996 water year.
[h] Flux estimated from SSCs based on a regression-derived rating curve calibrated using daily mean values for SSC and discharge collected during the 1997 water year.
[i] Flux estimated from SSCs based on a regression-derived rating curve calibrated using daily mean values for SSC and discharge collected during the 1998 water year.
[j] Flux estimated from SSCs based on a regression-derived rating curve calibrated using daily mean values for SSC and discharge collected during the 1999 water year.
[k] Flux estimated from SSCs based on a regression-derived rating curve calibrated using daily mean values for SSC and discharge collected during the 2000 water year.
[l] Flux estimated from SSCs based on a regression-derived rating curve calibrated using daily mean values for SSC and discharge covering the 15-year period from the 1981 through and including the 1995 water years, $n = 4881$.
[m] Flux estimated from SSCs based on a regression-derived rating curve calibrated using daily mean values for SSC and discharge covering the 20-year period from 1981 through and including the 2000 water years, $n = 6599$.

the revised NASQAN programme produced estimates of suspended sediment flux for the 5-year period (1996–2000 WY) ranging from 423 (+3%) Mt to 450 (+9%) Mt (Table II). The LOADEST model, calibrated with 72 data points collected over the 5-year period generated an estimate of 450 Mt, a 9% overestimate (Aulenbach B.T., USGS, personal communication, 2002), whereas a rating curve calibrated with the same data generated an estimate of 439 Mt, a 6% overestimate (Table II). On the other hand, the sediment rating curve generated from the daily SSCs for the same dates as the NASQAN samples, but collected by a different group under the auspices of another programme, generated a 5-year estimate of 447 Mt (Table II). Note that all these 5-year estimates display relatively small but significant differences between each other, and in comparison to the actual flux of 414 Mt. At least some of these differences probably can be ascribed to short-term spatial and temporal variations in SSC during the collection of actual samples and/or to laboratory imprecision (e.g. Horowitz, 1995; Horowitz *et al.*, 2001). Further, the positive bias for all these estimates probably can be ascribed to the NASQAN sampling schedule; although the sampling is hydrologically based in an attempt to cover at least 85% of the annual flow conditions at a site, it tended to emphasize non-baseflow periods (Hooper *et al.*, 2001). In turn, many of these relatively small but significant differences probably should be viewed as 'noise' that simply cannot be eliminated and which is inherent to the sediment rating-curve approach for determining daily SSCs for subsequent flux estimates.

When sediment rating curves were generated from the daily mean discharges and SSCs for each year from 1996 to 2000, and then used to estimate total flux for the entire 5-year period, the differences between the estimated and the actual fluxes increased substantially ($\pm$11%, Table II). On the other hand, the range of estimates (369–461 Mt) was almost evenly distributed around the actual flux of 414 Mt. This would imply that no one water year, between 1996 and 2000, sufficiently covered the variety of flow and concentration conditions extant during the entire 5-year period, to produce a more accurate estimate. The exception appears to be 1996, where the estimated flux was 409 Mt (Table II). However, a detailed examination of the percentile distributions of discharge and SSCs by year, failed to indicate why/how 1996 differed substantially from the other 4 years. Hence, the reason for the relatively good agreement between the actual flux, and the estimate based solely on 1996 data, relative to the fluxes based on data from each of the other 4 years, remains unresolved.

When sediment rating curves were generated using data from as long ago as 20 years (starting with the 1981 WY), the errors ($\leq$ 24%) in the 5-year flux estimate increased substantially, and display a marked positive bias relative to the actual flux (Table II). Based on the decline in positive bias between the flux estimates for a 15-year calibration (1981–1995 WY, 512 Mt, +24%) versus a 20-year calibration (1981–2000 WY, 485 Mt, +17%), it appears that the flow and SSC conditions at the Thebes site were substantially different before and after the 1996 WY. An examination of a plot of the actual versus the estimated flux for the 20-year period covering the 1981 to 2000 WY makes this difference visually obvious, and indicates that the onset of the change in the relationship between discharge and SSC can be dated to 1993 (Figure 3). Prior to 1993, the actual fluxes almost invariably exceed the estimated fluxes whereas after 1993, the reverse is true, and the estimates almost invariably exceed the actual fluxes (Figure 3).

Interestingly, the 1993 WY marked the last major flood in the Mississippi River Basin (Holmes, 1996). An examination of the median discharge and SSC levels at the Thebes site pre- and post-1993 clearly indicates the source of the change. Prior to 1993, the median discharge and SSC were about 5300 m^3 s^{-1} and 304 mg L^{-1} respectively; however after 1993, the median discharge and SSC were about 5800 m^3 s^{-1} and 206 mg L^{-1} respectively. These data indicate that whereas after 1993 the median discharge displayed a relative increase of 9%, the median SSC displayed a relative decline of 38%. Even when the data are adjusted for differences in annual discharge by comparing median flow-weighted SSCs, the relative decline was 34% (520 mg L^{-1} to 370 mg L^{-1}). Actual measurements in the general area around the Thebes site indicated as much as 4 m of channel-bed scour during the rise, and although aggradation was noted during the recession, it appears that erosion far exceeded deposition during and just after the 1993 flood (Holmes, 1996). Clearly, the flood removed substantial amounts of 'stored' bed sediment, eliminating a major source of SSC for the Thebes

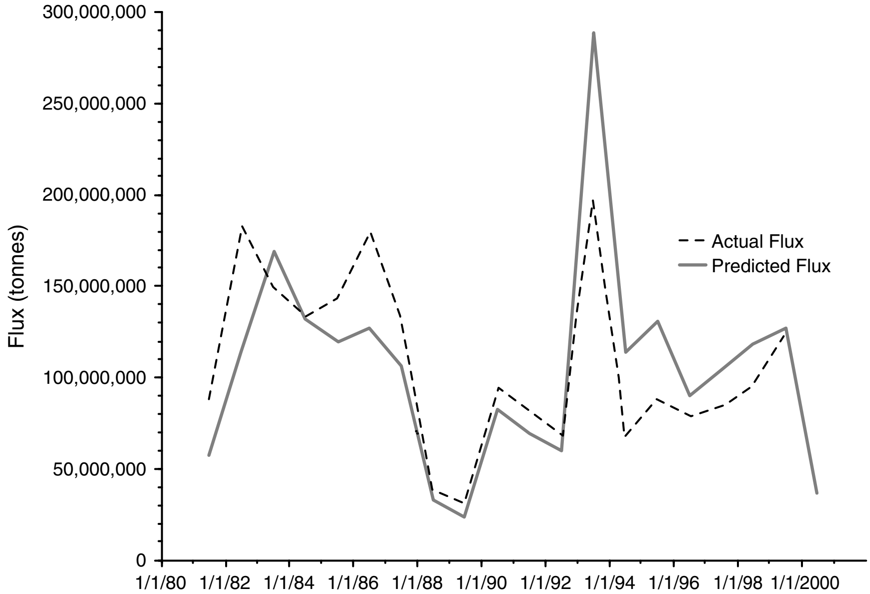


Figure 3. A comparison of actual and estimated annual fluxes (water years 1981–2000) for the Mississippi River at Thebes; the estimated fluxes are based on a single rating curve generated from daily mean discharge and daily SSC data

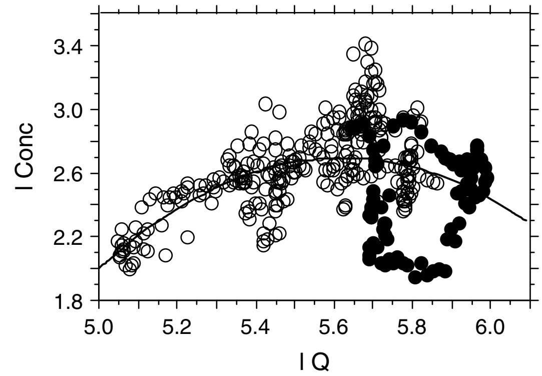


Figure 4. The 1993 water year sediment rating curve for the Mississippi River at Thebes; the data from samples collected during the flood are represented by the solid circles. Note that the flood samples appear to display a clockwise hysteresis loop. The labels on the graphs stand for log of discharge (l Q) and log of concentration (l Conc)

site. The rating curve for the 1993 water year, which is sharply convex, clearly supports the view that the Mississippi River at Thebes was 'sediment-starved' as a result of the flood (Figure 4).

An examination of the pre- and post-1993 sediment rating curves and the associated 5-year flux estimates for the other NASQAN Mississippi River mainstem sites, as well as those for the outlets from both the Missouri (Hermann) and Ohio Rivers (Grand Chain), indicates that the flood affected almost the entire basin (Table III). Note the substantial differences in 5-year suspended sediment flux estimates between sediment rating curves based on both data sets. Almost all the 1996 to 2000 WY site-specific cumulative suspended sediment flux estimates based on the pre-1993 data are substantially greater than those estimates based on the post-1993 data (Table III). Hence, the comparisons at Thebes, as well as for the other Mississippi River Basin sites, would appear to indicate that fluxes based on pre-1993 data probably represent overestimates.

Table III. The effect of using pre- and post-1993 SSC and discharge data to calibrate rating curves for predicting suspended sediment fluxes for selected Mississippi River Basin NASQAN sites

Site	Old data[a] *n*	New data[b] *n*	Flux (Mt) for water years 1996–2000 using:		
			All data	Old data	New data
Mississippi River at Clinton	128	62	20	21	19
Missouri River at Hermann	212	75	481	554	338
Mississippi River at Thebes	4151	2448	485	593	370
Ohio River near Grand Chain	221	65	144	153	124
Mississippi River at St Francisville	169	74	535	592	407

[a] Discharge and SSC data prior to 1993.
[b] Discharge and SSC data post 1993.

Even though the flood was confined to the upper part of the Mississippi River, the loss of stored sediment upstream almost certainly affected the lower part of the river as well (e.g. St Francisville, Table III). This probably was further exacerbated by the loss of locally stored bed sediment because even though the lower part of the river never reached flood stage in 1993, annual mean discharge (20 600 m^3 s^{-1}) at St Francisville was the second highest on record for the 1931 to 2000 WY period (C. Demas, USGS, personal communication, 2002).

The exception to this pattern of reduced post-1993 estimated suspended sediment fluxes occurred in the upper part of the Mississippi River (e.g. Clinton, Table III). The relatively unchanged level of estimated annual suspended sediment fluxes occurred even though the upper part of the river clearly exceeded flood stage in 1993, and experienced significant amounts of bed scour (Holmes, 1996). However, substantial downstream dispersion of the scoured material was limited by the extensive lock and dam system in the upper part of the Mississippi River (R. Meade, USGS, personal communication, 2002). Hence, the bed sediment source, although somewhat displaced, never really left the area.

Temporal resolution and associated errors for 5-year sediment rating-curve-based flux estimates

Because the Mississippi River at Thebes is a daily sediment station, it provides an ideal case for examining both the nature of the errors associated with the use of the sediment rating-curve method of estimating SSCs (fluxes), as well as the relative errors associated with different levels of temporal resolution. Examination of a plot of the actual versus the calculated daily fluxes for the site for the 5-year 1996–2000 WY period indicates that the sediment rating-curve method tends to underpredict the highs and overpredict the lows, even when daily discharge and SSC data are used (Figure 5a). In other words, the method fails to encompass the complete range of variance in the daily fluxes (SSCs) at the site. Typically, there are relatively few highs (events) compared with a large number of lows (baseflow), and the underpredictions are substantially larger than the overpredictions. As such, the range of errors associated with relatively short time-frames (e.g. daily, weekly) are likely to be substantially larger than those associated with longer time-frames because the over- and underpredictions do not have sufficient time to balance each other (Table IV and Figures 5a–5e). Hence, for the Mississippi River at Thebes, where 5 years of daily SSC and discharge data were used to develop the sediment rating curve, the percentage differences range from −76 to +205% for daily values, from −63 to +157% for weekly values, from −43 to +58% for monthly values, −33 to +21% for quarterly values, −12 to +8% for annual values, and <1% for 5 or more years (Table IV). The reduction in the percentage differences between actual and predicted suspended sediment fluxes with decreasing temporal resolution (longer time-frames), is even more apparent when examining the absolute values for the means/medians of the percentage differences between the two, which range from a high of 26% (mean)/18% (median) for daily estimates, to a

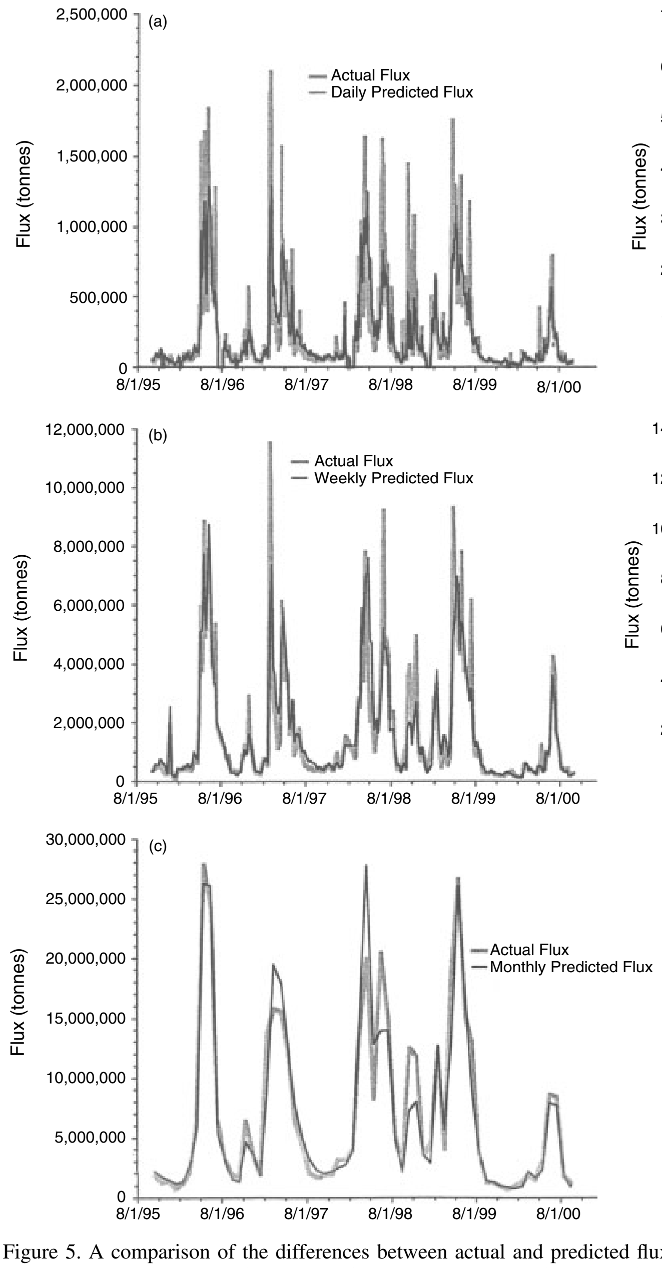


Figure 5. A comparison of the differences between actual and predicted fluxes for various levels of temporal resolution for the Mississippi River at Thebes during the 1996–2000 water year period: (a) daily, (b) weekly, (c) monthly, (d) quarterly and (e) annually

low of 8%/8% for annual estimates (Table IV). In fact, the percentage difference declines to <1%/<1% for the entire 5-year period.

The patterns of prediction for sediment rating curve-derived daily SSCs (fluxes) displayed for the Mississippi River at Thebes site are mirrored in all the other NASQAN basins and sampling sites where only discontinuous data are available (Horowitz *et al*., 2001). That is, all the sediment rating curves tend to underpredict the high and overpredict the low daily SSCs at their respective sites. This conclusion is based on a comparison between

Table IV. Various levels of temporal resolution and their associated errors for the Mississippi River at Thebes (water years 1996–2000) and for the Rhine River at Maxau (water years 1989–1993)

Site	Period	Maximum underestimate			Maximum overestimate			Absolute	
		Actual flux (t)	Estimated flux (t)	Percentage difference	Actual flux (t)	Estimated flux (t)	Percentage difference	Mean percentage difference	Median percentage difference
Mississippi River at Thebes (water years 1996–2000)	Daily	500 000[a]	110 000	−76	400 000	1 100 000	205	26	18
	Weekly	1 300 000	480 000	−63	3 000 000	7 600 000	157	24	18
	Monthly	13 000 000	7 200 000	−43	8 300 000	13 000 000	58	19	15
	Quarterly	28 000 000	19 000 000	−33	8 200 000	9 900 000	21	13	12
	Annually	32 000 000	28 000 000	−12	96 000 000	100 000 000	8	8	8
Rhine River at Maxau (water years 1989–1993)	Daily	22 000	5 700	−74	160	1 600	946	44	27
	Weekly	27 000	12 000	−56	1 800	9 600	434	33	23
	Monthly	70 000	45 000	−36	12 000	45 000	274	28	22
	Quarterly	240 000	160 000	−34	200 000	300 000	54	19	17
	Annually	1 100 000	1 000 000	−8	740 000	940 000	27	10	7

[a] Fluxes are rounded to two significant figures, but percent differences are based on the actual calculated values.

the actual sample SSCs from the various site-specific calibration sets and the matching sediment rating-curve-derived predicted values (Horowitz *et al*., 2001). On that comparative basis, the errors associated with annual sediment rating-curve-derived suspended sediment flux estimates ranged from -23% to $+20\%$, but typically were $\leq 15\%$. It should be noted that within the NASQAN programme, the errors for sediment rating-curve-derived annual fluxes tended to be larger for sand-dominated rivers than those where the suspended sediment was predominantly silt/clay; hence, this approach for estimating SSCs and fluxes may be less effective in the former than in the latter.

A similar set of predictions, calculations and comparisons for the Rhine River at Maxau, for a different 5-year period (1989–1993 WY), again using daily discharge and SSC data, indicate that the patterns observed for the Mississippi River at Thebes site, as well as those for the other NASQAN sites, are not atypical, at least for medium to large rivers (Table IV and Figure 6). Hence, the Rhine sediment rating curves tend to underpredict the high and overpredict the low daily SSCs, and estimation errors decline with decreasing temporal resolution. Note that the larger overestimation errors associated with the Rhine SSC and flux predictions (Table IV) typically occur at very low SSCs (< 20 mg L^{-1}), where actual measurement errors also are greatest (e.g. Horowitz, 1995). Such relatively low SSCs tend to be absent at most of the NASQAN sites, except in the Columbia River Basin (Horowitz *et al*., 2001).

From the foregoing, it should be obvious that the size of the errors associated with rating-curve-derived suspended sediment fluxes increase with increasing temporal resolution. This may be of particular concern if this method of flux estimation is applied to the determination of TMDLs (total maximum daily loads), especially if the 'D' (daily) in TMDLs is taken literally. This is a particularly contentious issue in the USA (e.g. Keyes and Radcliffe, 2002). Regardless, end-users of this type of rating-curve-derived flux data must decide at which point the errors associated with various levels of temporal resolution become unacceptably large and/or are too large to facilitate sound management decisions.

The differences between actual and sediment rating-curve-derived flux calculations/estimations should be evaluated in light of the potential errors associated with the data used to derive them. Errors associated with discharge measurements typically can range from between 2 to 20%, depending on the method used and the conditions extant at the time of the measurement, however, mean errors, at the 95% confidence limit, are of the order of ± 5–8% (Kennedy, 1983; Sauer and Meyer, 1992). The errors associated with SSC measurements tend to vary with concentration and grain-size distribution, as well as the site conditions at the time of collection. However, based on replicate samples, at concentrations above 20 mg L^{-1}, it is not atypical to see differences

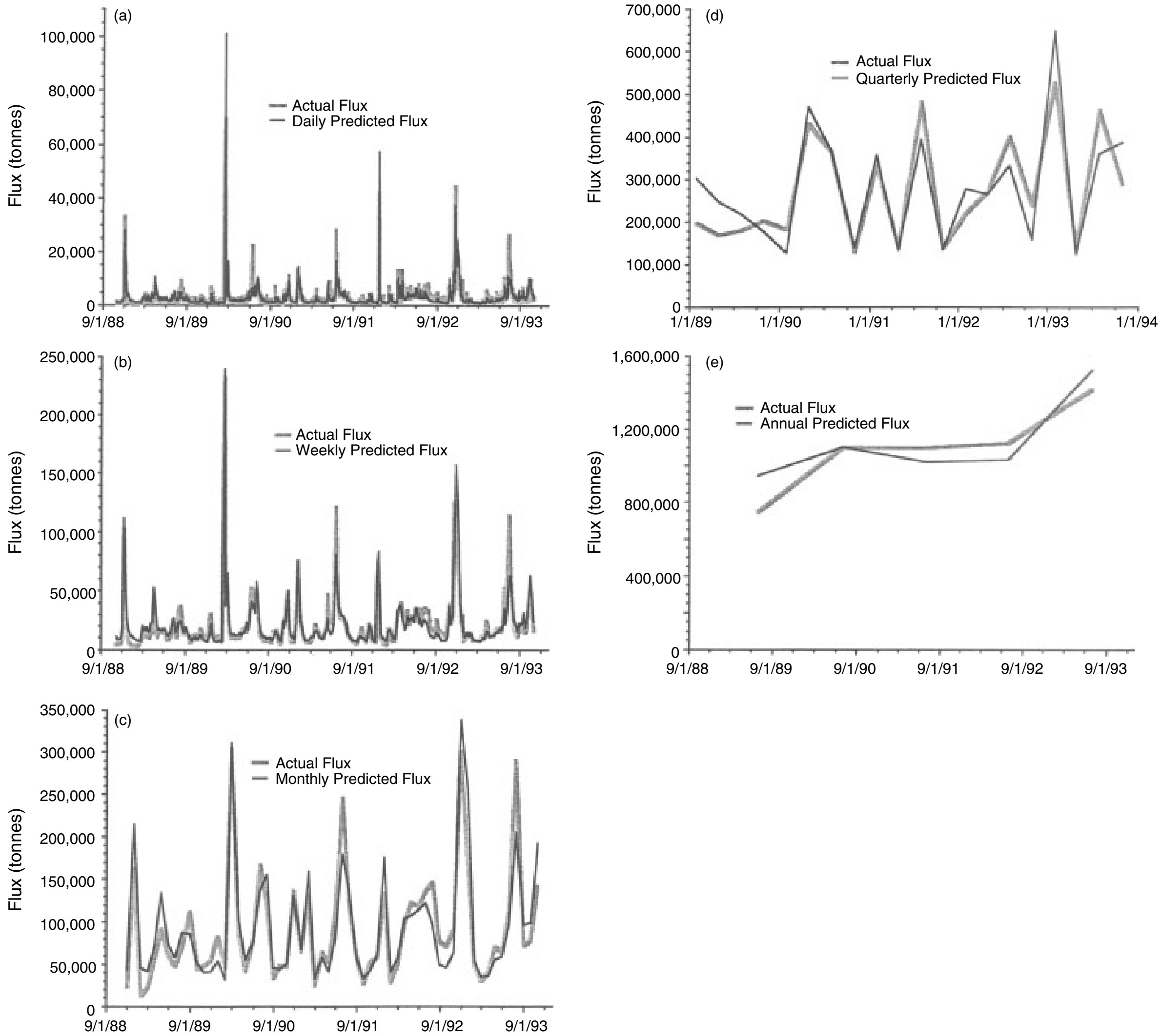


Figure 6. A comparison of the differences between actual and predicted fluxes for various levels of temporal resolution for the Rhine River at Maxau during the 1989–1993 water year period: (a) daily, (b) weekly, (c) monthly, (d) quarterly and (e) annually

of the order of ±10% (Edwards and Glysson, 1988; Horowitz *et al*., 2001). These replicate SSC differences result from a combination of factors including, but not limited to, naturally occurring short-term spatial and temporal variability in cross-sectional SSC (e.g. Horowitz, 1995). Finally, at least based on the current work, for temporal resolutions of a year or longer, there appears to be some degree of inherent 'noise' (of the order of ±5%) in the rating-curve approach as well. As a result of the foregoing, differences between actual and rating-curve-derived flux estimates (predictions) of the order of ±15–20% should be viewed as falling within the normal range of measurement/prediction error.

Temporal resolution and associated errors for 20-year rating-curve-based flux estimates and the issue of interannual variability

The actual 20-year cumulative suspended sediment flux for the Mississippi River at Thebes site for the 1981–2000 WY period is 2100 Mt (Table V). A single sediment rating curve using the entire 20-year data set (a corrected second-order polynomial, Figure 7a), yielded an estimate of 2000 Mt for the same period, only a 5% underestimate (Table V). This is a fairly standard approach for generating site-specific sediment rating curves where long-term data are available, and is predicated on the assumption that all the data from the site are part of the same statistical population. Note that despite the close agreement between the actual and estimated cumulative suspended sediment flux for the 20-year period using the single rating-curve approach, the differences ($\leq \pm 48\%$ with an absolute mean of 17%) between the actual and estimated fluxes for individual years (e.g. 1982, 1986, 1989, 1993 and 1995), can be substantial (Table V and Figure 7a). On the other hand, when individual annual sediment rating curves for the Thebes site are generated for the same 20-year period, and the associated annual suspended sediment fluxes summed, only a marginal improvement is observed in the estimated cumulative flux (Table V and Figure 7b). The sum of the estimated annual suspended sediment fluxes of 2100 Mt matches the actual flux, within rounding errors (Table V). However, the individual annual suspended sediment flux estimates are now significantly closer ($\leq \pm 8\%$ with an absolute mean of 2%) to the actual annual fluxes than could be achieved with the single rating-curve approach (Table V and Figure 7b).

An examination of the records and sediment rating curves for a number of other long-term daily sediment monitoring sites indicates that these conclusions are not confined to the Mississippi River at Thebes. Similar patterns of marginally improved long-term flux estimates, and substantially improved annual estimates, occur in such diverse locations as the Rhine River at Maxau, Germany, the Yadkin River at Yadkin College, North Carolina, the Green River at Munford, Kentucky and the Schuykill River at Berne, Pennsylvania (Figure 8). The annual median suspended sediment flux for the Rhine River at the Maxau site for the 20-year (1974–1993 WY) period was 1·15 Mt based on the results from individually generated annual sediment rating curves. This is on a par with the median load (1·16 Mt) reported by Asselman (2000) for the slightly shorter 1974 to 1990 WY period. The Thebes site represents a drainage area in excess of 1 800 000 km^2; however, these additional basins range downward to $<$1000 km^2 (Schuykill; Table I). Hence, the conclusions regarding improved predictions would appear to apply to medium and relatively small rivers as well. However, note that the improvements in annual suspended sediment flux estimates tend to decline with decreasing basin size (Figure 8). This probably is a reflection of the 'flashier' nature of, and/or that the annual interrelations between discharge and SSC are less consistent/display greater variability in, smaller rivers relative to larger ones. In the case of really small watersheds, which can display marked short-term variability in discharge and SSC, the single and/or annual rating-curve approach may be ineffective, even at temporal resolutions of a year or more (e.g. Walling, 1977).

The individual annual sediment rating curves for the Thebes site appear to provide some insights into the differences between the estimates generated by the two approaches, cumulative flux from a single rating curve versus cumulative flux from individual annual sediment rating curves for the same time period (Figure 9). It would appear that the 20-year data set either does not consist of a single statistical population and/or the interrelation between discharge and SSC has not been constant for the entire period. The latter point already has been cited in terms of the pre- and post-1993 differences in median discharge and SSC, as well as the pre- and post-1993 differences between actual and estimated annual fluxes. However, even within the pre- and post-1993 periods, there are marked differences between annual sediment rating curves (Figure 9). Some curves are based on linear functions (e.g. 1982, 1984, 1999), whereas the majority are based on second-order polynomial functions [either concave (e.g. 1990, 1997, 2000) or convex (e.g. 1983, 1987, 1993)].

Several authors have attempted to relate the shapes of sediment rating curves to specific hydrological factors/processes (e.g. Sichingabula, 1998; Asselman, 2000). The shape variations (linear, convex, concave) associated with the annual sediment rating curves for the Thebes site do appear to reflect either the relative flow conditions for that year and/or a prior event (e.g. a flood). During the 1981–1992 period, median flow at the Mississippi River at Thebes site was about 5300 m^3 s^{-1}, whereas during the 1994–2000 WY period the median discharge was about 5800 m^3 s^{-1}; during the 1993 flood year, median discharge was 12 600 m^3 s^{-1}. As already noted, the major 1993 WY flood on the Mississippi River led to a severely convex rating curve, which indicates that the site probably was 'sediment-starved' as a result (Figure 9). Also note that the impact of the flood extended through the following 3 years (1994–1996 WY), which also had convex sediment rating

Table V. A comparison of actual and estimated cumulative suspended sediment fluxes for the Mississippi River at Thebes site based either on a 20-year rating curve or on annual rating curves

			Linear regression		Polynomial regression			
Rating curve period	Water year	Actual flux (Mt)	Predicted flux (Mt)	Corrected predicted flux (Mt)	Predicted flux (Mt)	Corrected predicted flux (Mt)	Best-fit predicted flux (Mt)	Percentage difference
Annual fluxes based on a 20-year rating curve	1981	88	49	56	50	57	57	−35
	1982	180	99	110	99	110	110	−39
	1983	150	160	180	150	170	150	<1
	1984	130	120	140	120	130	130	0
	1985	140	110	120	100	120	120	−14
	1986	180	110	130	110	130	130	−28
	1987	130	93	110	90	100	100	−23
	1988	39	28	32	29	33	33	−15
	1989	30	20	23	20	23	23	−23
	1990	94	72	82	70	80	80	−15
	1991	81	60	68	61	69	69	−15
	1992	68	51	58	53	60	60	−12
	1993	200	290	330	250	290	250	25
	1994	67	100	120	99	110	99	48
	1995	89	120	140	120	130	120	35
	1996	79	82	94	79	90	79	<1
	1997	84	92	110	91	100	91	8
	1998	96	110	120	100	120	100	4
	1999	120	110	130	110	130	130	8
	2000	32	32	37	33	37	32	<1
	Sum	2100					2000	
	% difference						−5	
Annual fluxes based on annual rating curves	1981	88	85	93	82	91	91	3
	1982	180	170	190	170	190	190	6
	1983	150	140	160	140	150	150	<1
	1984	130	120	140	130	140	140	8
	1985	140	130	140	130	140	140	<1
	1986	180	160	180	160	180	180	<1
	1987	130	120	140	120	130	130	<1
	1988	39	38	41	37	39	39	<1
	1989	30	24	28	25	29	29	−3
	1990	94	80	89	89	98	98	4
	1991	81	72	79	77	83	83	2
	1992	68	60	65	62	67	67	−1
	1993	200	170	210	170	200	200	<1
	1994	67	65	71	61	67	67	<1
	1995	89	91	96	84	88	88	−1
	1996	79	77	81	75	79	79	<1
	1997	84	75	80	78	83	83	−1
	1998	96	89	95	88	94	95	−1
	1999	120	120	120	120	120	120	<1
	2000	32	30	31	31	32	32	<1
	Sum	2100					2100	
	% difference						<1	

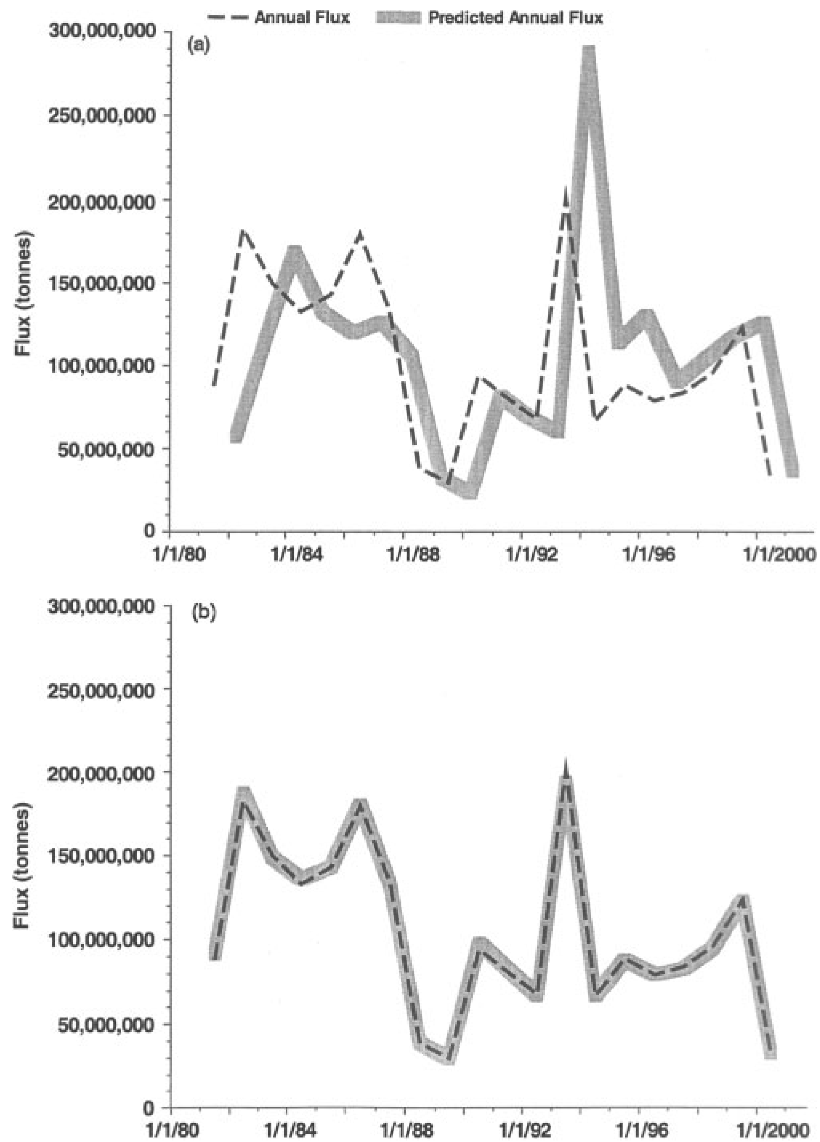


Figure 7. A comparison of actual annual fluxes for the Mississippi River at Thebes site for the 1981–2000 water year period versus estimated fluxes calculated from (a) a single 20-year rating curve and from (b) individual annual rating curves

curves; albeit, the curve for the 1996 WY is barely so (Figure 9). This may indicate that it took about 3 years for local conditions to recover from the effects of the 1993 flood. Additional convex sediment rating curves occurred in 1983 and 1987; the former also was a very high flow year (median discharge was 9900 m^3 s^{-1}). In this case, the site appeared to recover rather rapidly as the curve for 1984 was linear (Figure 9). Lastly, the 1987 rating curve also was convex, although the flow (7100 m^3 s^{-1}) for that year just about matched the median for the period. This appears to be an artifact of the period designating a water year (1 October to 30 September). The high flows leading to the loss of locally stored sediment occurred during the last months of the 1986 WY. Note that the rating curve for 1986 also is slightly convex, as is the one for 1988, probably indicating that it took more than a year for the site to recover (Figure 9).

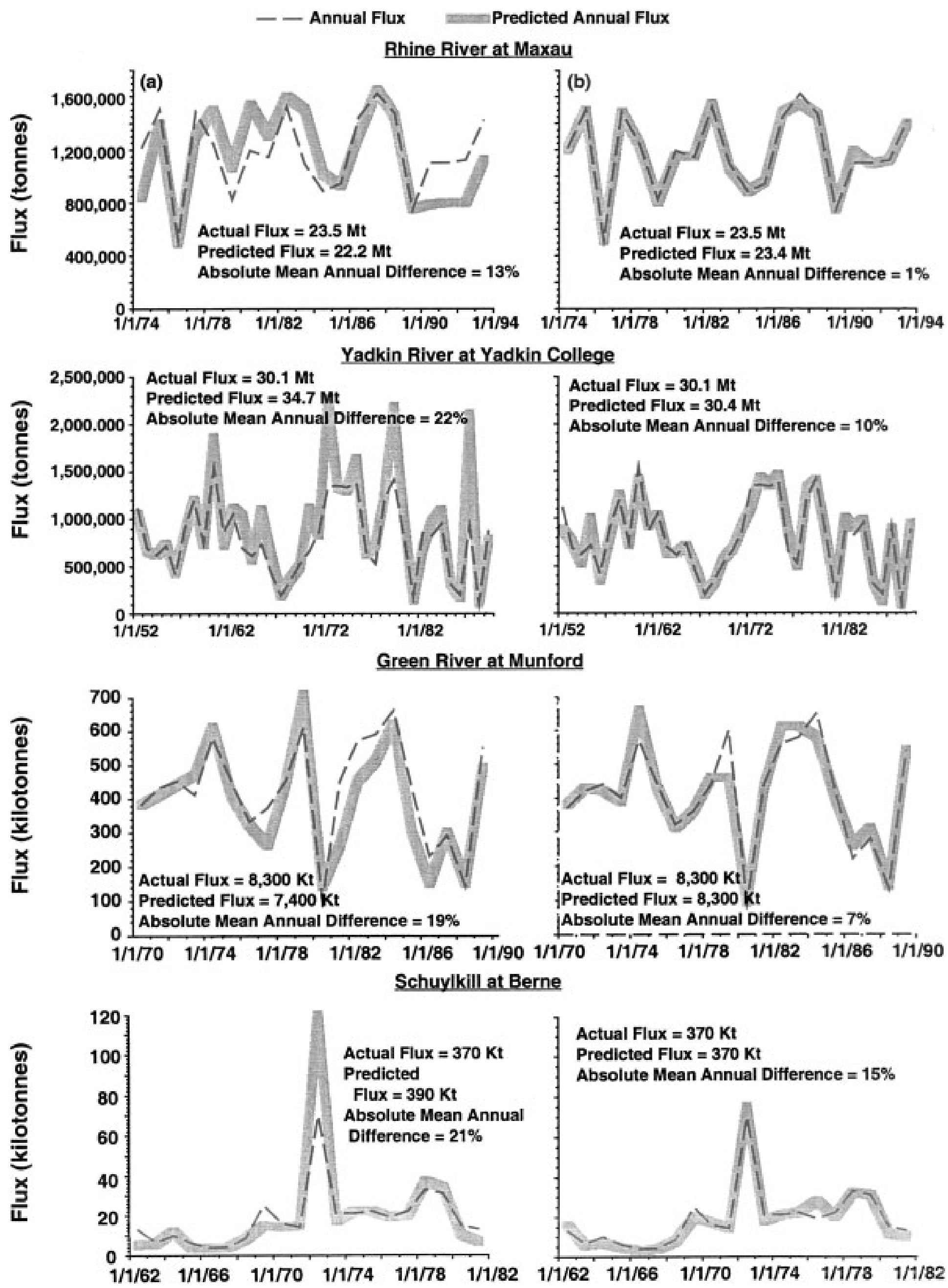


Figure 8. Comparisons of actual and estimated annual fluxes for selected rivers; the estimated fluxes were calculated from: (a) a single rating curve covering the entire period of interest and from (b) individual annual rating curves

On the other hand, concave sediment rating curves at the Thebes site occurred during the 1989 to 1992 WYs, as well as in the 1997 and 2000 WYs (Figure 9). Most of these appear to be related to periods of lower than median flow (e.g. 1989–1992 WYs (3700 $m^3\ s^{-1}$ to 6200 $m^3\ s^{-1}$) and 2000 WY(4400 $m^3\ s^{-1}$)). On the other hand, the 1997 concave curve occurred during a median flow year (7600 $m^3\ s^{-1}$). The concave nature of these

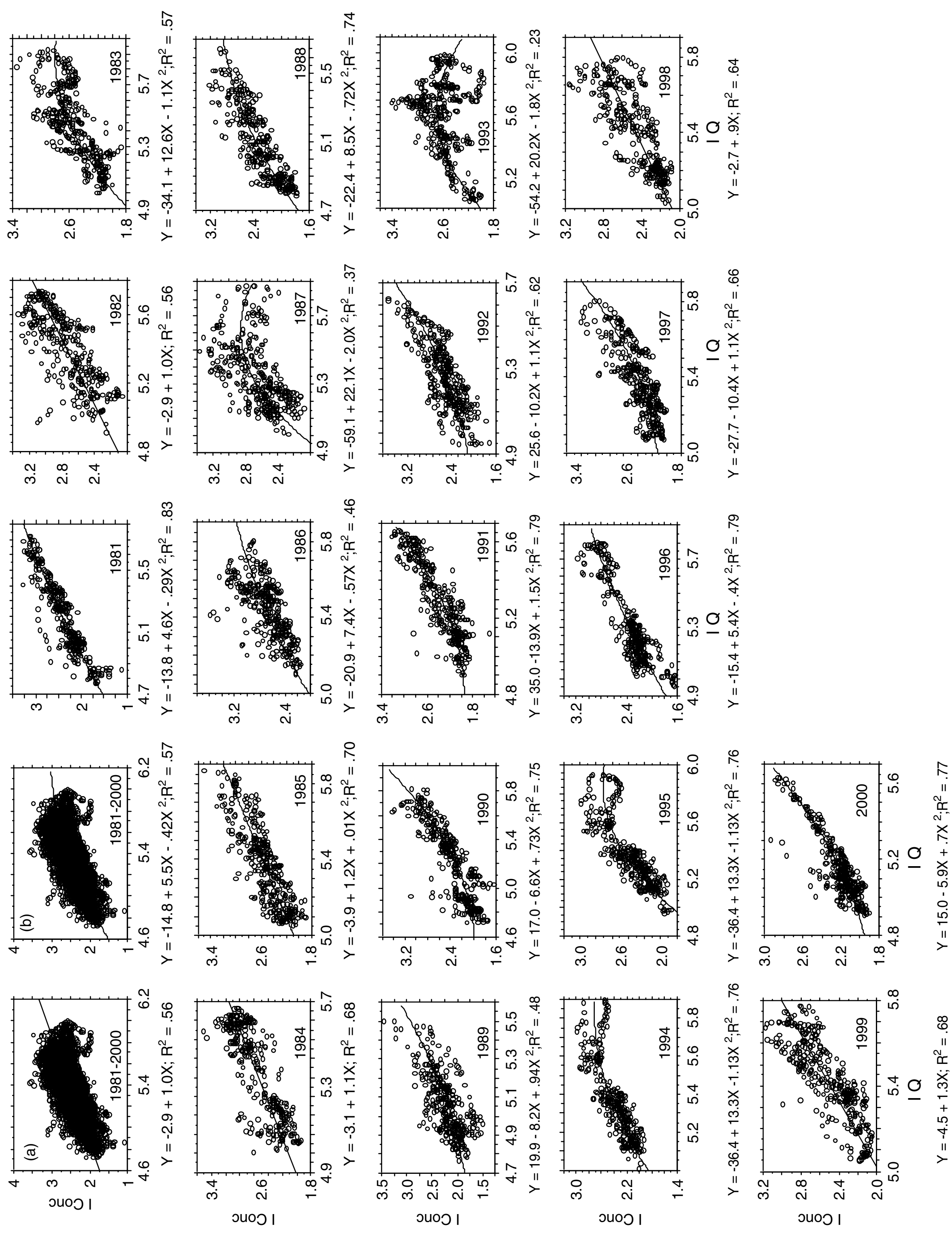


Figure 9. Twenty-year and annual sediment rating curves for the Mississippi River at Thebes site covering the 1981–2000 water year period; the 20-year sediment rating curves are (a) linear and (b) second-order polynomial regressions. The labels on the graphs stand for log of discharge (l Q) and log of concentration (l Conc)

curves may be the result of resuspension, and potential downstream transport of locally stored bed sediment, coincident with higher flows occurring during these periods. These observations tend to support the view that the shape of an annual sediment rating curve may be used to infer various hydrological processes/conditions and/or to determine how long it may take for a site to recover from the effects of a major disruption (e.g. a flood); however, this conclusion should be applied with some degree of caution, at least until further evaluations are completed.

One interesting statistical observation can be made from examining all the sediment rating curves in Figure 9. The r^2 statistic normally is used as an indicator of the so-called 'goodness of fit' for regression equations. However, when it comes to the accuracy of rating-curve-derived annual flux estimates, it appears to be rather insensitive. For example, the r^2 value for the 1993 WY rating curve was 0·23, yet the predicted flux was 195 Mt versus an actual flux of 201 Mt (both appear as 200 Mt in the table as a result of rounding), a 5% underestimate (Figure 9 and Table V). On the other hand, the r^2 for 1981 is 0·83, yet the difference between the predicted flux of 91 Mt versus an actual flux of 88 Mt (+3%) is not substantially better, in terms of magnitude, than the difference for 1993 (Figure 9 and Table V). Hence, at least in the case of sediment rating curves, a low r^2 statistic may be a poor indicator of the accuracy of the subsequent annual flux estimates. In fact, the key to a good rating-curve-derived flux estimate appears to be how well the regression averages out the 'scatter' in the data, rather than how well the curve actually fits all the data points.

The impact of sampling frequency on the accuracy of rating-curve-derived annual and 5-year flux estimates

As a result of resource (monetary/personnel) constraints, most monitoring programme designs tend to represent compromises in which sampling frequency is balanced against a presumptive requirement for a certain level of accuracy or an acceptable level of error. The underlying issue is what levels of sampling frequency are required to achieve a minimally acceptable level of measurement/predictive accuracy? Stated another way, how much error can be tolerated while still permitting sound management decisions relative to the problem at hand? Data from long-term daily monitoring sites can be used to address these questions relative to sediment fluxes.

Based on the previous results, which indicated that the effective temporal resolution of rating-curve-derived flux estimates were quarterly or greater, this evaluation was limited to periods of 1 and 5 years. There were two potential methods for selecting the various data sets for this exercise: (i) random selection, as occurs in Monte Carlo type simulations (e.g. de Vries and Klavers, 1994); or (ii) calendar based selection. The latter approach was used for both the 1- and 5-year evaluations because it better represents the method of sample scheduling used in the majority of monitoring programmes. Hence, for an evaluation of samples collected on a 15-day schedule, for a 1-year period, each day of the year was assigned a number from 1 to 15 consecutively, beginning with 1 October and ending with 30 September. Then, all the data from the days having the same number were combined into 15 separate data sets, and a separate rating curve was generated for each set. In turn, each rating curve was used to generate SSC data for the missing days and annual flux estimates were calculated for each of the 15 rating curves (Figure 10). Thus, for example, the 10-day annual evaluations generated 10 rating-curve-derived annual flux estimates, the 30-day annual evaluations generated 30 rating-curve-derived annual flux estimates, the 50-day 5-year evaluations generated 50 rating-curve-derived 5-year estimates, etc.

The effect of sampling frequency on the accuracy and associated errors of 5-year rating-curve-derived suspended sediment flux estimates entailed using the daily discharge and SSC values for the Mississippi River at Thebes site for the 1996–2000 WY period and the Rhine at Maxau site for the 1989–1993 WY period. The sampling frequencies evaluated in this way corresponded to: (i) once a day; (ii) once every other day; (iii) once every 3 days; (iv) once every 4 days; (v) once every 5 days (weekly); (vi) once every 10 days, (vii) once every 25 days (monthly); and (viii) once every 50 days (every other month).

Not surprisingly the accuracy of the 5-year estimates decreased, and the size and range of the associated errors increased with decreasing sampling frequency (Figure 10). Also note that the median 5-year flux values

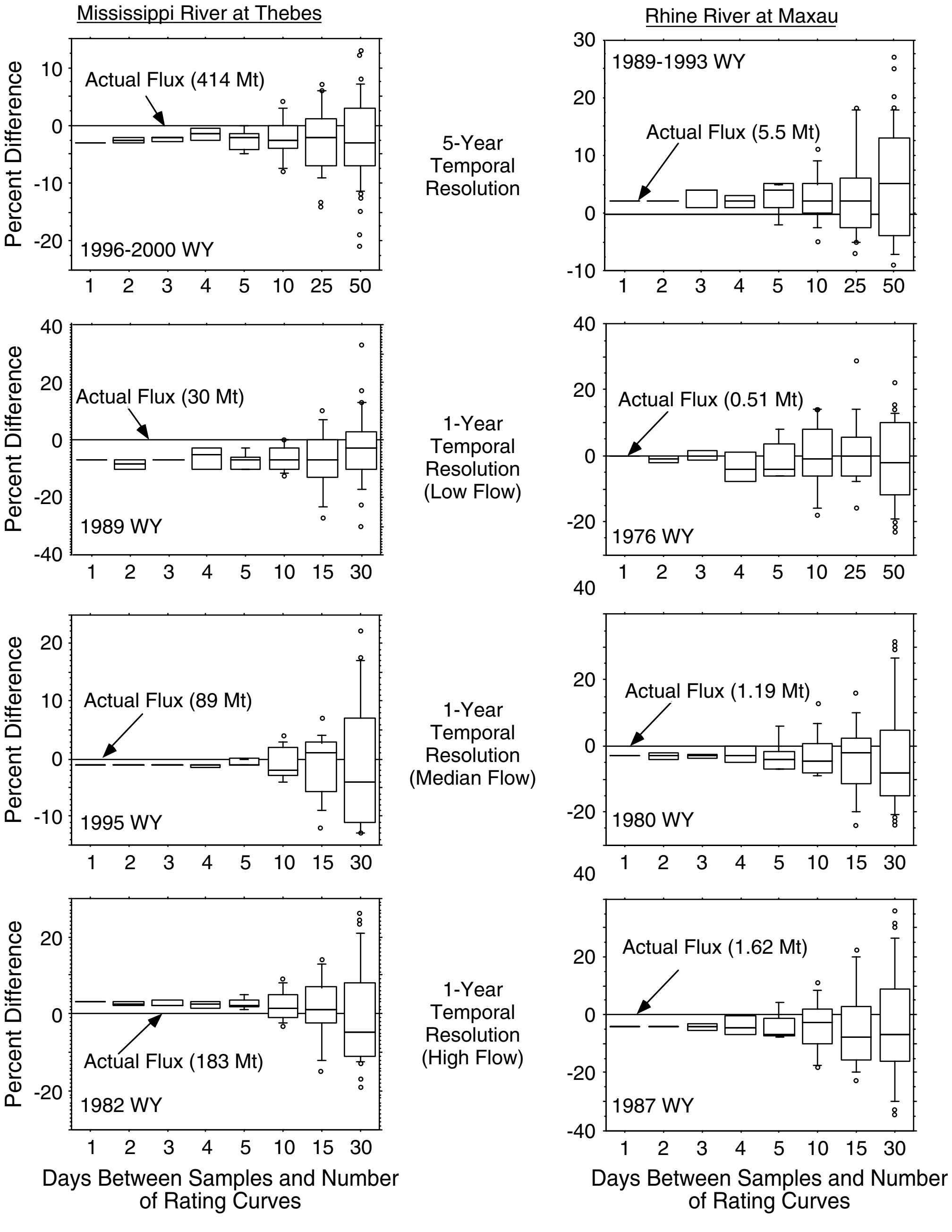


Figure 10. The effect of sampling frequency on the estimation of suspended sediment fluxes for the Mississippi River at Thebes and the Rhine at Maxau sites covering a 5-year period and three 1-year periods. The single year plots cover high, median and low flux years

for all the sampling frequencies evaluated for the Thebes site displayed a small but significant negative bias, whereas the Maxau site displayed a small but significant positive bias, even though a 'smearing' correction was applied to second-order polynomial sediment rating curves in both cases. Despite this bias difference, both sets of calculations produced similar results relative to sampling frequency. There was little difference between sampling frequencies ranging from 1 to 5 days. On the other hand, estimation errors from sampling frequencies of the order of once every 2 months (once every 50 days) were little compromised, and tended to fall within a range of $\leq\pm 20\%$. Because the calculations were based solely on calendar distributions, they probably represent the maximum error likely to occur with this level of sampling frequency (Figure 10). If the same level of sampling (once every 50 days) were hydrologically distributed, such as to encompass some 80 to 85% of the typical range of discharge at these sites, the associated estimation errors likely would be substantially less (e.g. Horowitz, 1995).

The effect of sampling frequency on the accuracy and associated errors of annual suspended sediment flux estimates for both Thebes and Maxau also was investigated concurrently (Figure 10). These evaluations covered high (Thebes, 1982; Maxau, 1987), median (Thebes, 1995; Maxau, 1980), and low (Thebes, 1989; Maxau, 1976) flux years. The sampling frequencies evaluated in this way corresponded to: (i) once a day; (ii) once every other day; (iii) once every 3 days; (iv) once every 4 days; (v) once every 5 days (weekly); (vi) once every 10 days, (vii) once every 15 days (fortnightly); and (viii) once every 30 days (monthly). The sediment rating curves for all of the Thebes and the high and median flux years for Maxau were 'smearing' corrected second-order polynomials. However, the sediment rating curves for the Maxau low flux year were 'smearing' corrected linear ones. Note that as with the 5-year study, there is little difference between 1- and 5-day sampling frequencies. Further, even collecting a sample as infrequently as once a month produced differences only of the order of $\leq\pm 20\%$, regardless of the flux levels (high, low or median). The same caveats apply to the annual, as to the 5-year study, hence, hydrologically based sampling, as opposed to calendar-based sampling, is likely to produce substantially more accurate estimates.

CONCLUSIONS

1. The composition of calibration sets can have a significant impact on suspended sediment rating curves. It is not axiomatic that long-term data sets from the same site, even when generated by consistent methods, represent a single statistical population. This assumption should be carefully evaluated prior to using all the data to produce a single sediment rating curve for the estimation of suspended sediment fluxes for periods of at least 5 years, or less.
2. Despite the foregoing, excellent cumulative flux estimates (errors of $<\pm 1\%$) covering periods of 20-years or more, can be generated using a single sediment rating curve based on data spanning the entire period; however, somewhat better estimates for the entire period, and markedly better annual estimates within the period can be obtained if individual annual sediment rating curves are used instead.
3. The 1993 Mississippi River flood substantially altered the interrelation between discharge and SSC throughout much of the river, as well as its main tributaries, for the period covering the 1981 to 2000 WYs.
4. Sediment rating curves tend to underpredict high, and overpredict low suspended sediment concentrations; as there typically are relatively few highs (events) compared with a large number of lows (baseflow), the errors associated with short time-frames (daily, weekly) tend to be larger than those associated with longer periods (quarterly, annually).
5. It appears that sediment rating curves can be used to generate reasonably accurate ($\leq$15–20%) suspended sediment flux estimates for quarterly time-frames or greater, for large, medium and relatively small rivers. As such, for these longer time-frames, sediment rating curves represent a markedly cheaper alternative than such methods as automatic samplers or turbidometers.
6. Some evidence exists to indicate that the shapes (linear, concave, convex) of annual sediment rating curves can be used to infer hydrological factors/processes.
7. Sampling frequencies can exercise a substantial impact on the accuracy of sediment rating-curve-generated flux estimates; however, good estimates (errors of $\leq\pm 20\%$) can be obtained from relatively infrequent

samples. This is particularly true if the samples are collected on a hydrological as opposed to a calendar-basis; good 5-year flux estimates appear to require as few as six samples per year, whereas good annual estimates appear to require as few as 12 samples per year.

ACKNOWLEDGEMENTS

I would particularly like to thank my many past and present USGS colleagues who have laboured for many years to collect long-term, daily, suspended sediment and discharge data under a wide variety of conditions. I also thank Charles Demas (USGS), Bob Meade (USGS), Michel Meybeck (University of Paris VI), Tim Stamey (USGS) and Des Walling (University of Exeter) for much useful information, discussion and their comments/suggestions.

REFERENCES

Asselman NEM. 2000. Fitting and interpretation of sediment rating curves. *Journal of Hydrology* **234**: 228–248.

Christensen VG. 2001. Characterization of surface-water quality based on real-time monitoring and regression analysis, Quivira National Wildlife Refuge, south-central Kansas, December 1998 through June 2001. *U.S. Geological Survey Water-resources Investigations* **01–4248**: 28 pp.

Crawford CG. 1996. *Estimating mean constituent loads in rivers by the rating-curve and flow duration rating-curve methods*. Unpublished PhD dissertation, Indiana University: Bloomington, IN; 245 pp.

De Vries A, Klavers HC. 1994. Riverine fluxes of pollutants: monitoring strategy first, calculation methods second. *European Journal of Water Pollution Control* **4**: 12–17.

DeutscherVerband für Wasserwirtschaft und Kulturbau (DVWK). 1986. *Regeln zur Wasserwirtschaft*, Schwedstoffmessungen. Regeln zur Wasserwirtschaft; 125 pp.

Duan N. 1983. Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Society* **78**: 605–610.

Edwards TK, Glysson GD. 1988. Field methods for measurement of fluvial sediment. *U.S. Geological Survey Open-file Report* **86–531**: 118 pp.

Ferguson RI. 1986. River loads underestimated by rating curves. *Water Resources Research* **22**: 74–76.

Hansen D, Bray DI. 1993. Single-station estimates of suspended sediment loads using sediment rating curves. *Canadian Journal of Civil Engineering* **20**: 133–143.

Holmes Jr RR. 1996. Sediment transport in the lower Missouri and the central Mississippi Rivers, June 26 through September 14, 1993. *U.S. Geological Survey Circular* **1120-I**: 23 pp.

Holtschlag DJ. 2001. Optimal estimation of suspended-sediment concentrations in streams. *Hydrological Processes* **15**: 1133–1156.

Hooper RP, Aulenbach BT, Kelly VJ. 2001. The National Stream Quality Accounting Network: a flux-based approach to monitoring the water of large rivers. *Hydrological Processes* **15**: 1089–1106.

Horowitz AJ. 1995. *The Use of Suspended Sediment and Associated Trace Elements in Water Quality Studies*. IAHS Special Publication No. 4, IAHS Press: Wallingford; 58 pp.

Horowitz AJ, Elrick KA, Smith JJ. 2001. Estimating suspended sediment and trace element fluxes in large river basins: methodological considerations as applied to the NASQAN programme. *Hydrological Processes* **15**: 1107–1132.

Kennedy EJ. 1983. Computation of continuous records of streamflow. In *Techniques of Water-resources Investigations of the U.S. Geological Survey*. Chapter A13, Book 3, U.S. Geological Survey: Washington, D.C. 53 pp.

Keyes AM, Radcliffe D. 2002. *A Protocol for Establishing Sediment TMDLs*. The Georgia Conservancy: Atlanta, GA; 31 pp.

Phillips JM, Webb BW, Walling DE, Leeks GJL. 1999. Estimating the suspended sediment loads of rivers in the LOIS study area using infrequent samples. *Hydrological Processes* **13**: 1035–1050.

Porterfield G. 1977. Computation of fluvial-sediment discharge. *Techniques of Water-resources Investigations of the U.S. Geological Survey*. Chapter C3, Book 3, U.S. Geological Survey: Washington, D.C. 66 pp.

Roberts G. 1997. The influence of sampling frequency on streamflow chemical loads. *Journal of the Chartered Institution of Water and Environmental Management* **11**: 114–118.

Sauer VB, Meyer RW. 1992. Determination of error in individual discharge measurements. *U.S. Geological Survey Open-file Report* **92–144**: 21 pp.

Sichingabula HM. 1998. Factors controlling variations in suspended sediment concentration for single-valued sediment rating curves, Fraser River, British Columbia, Canada. *Hydrological Processes* **13**: 1035–1050.

Walling DE. 1977. Assessing the accuracy of suspended sediment rating curves for a small basin. *Water Resources Research* **12**: 1869–1894.

Walling DE, Webb BW. 1981. The reliability of suspended sediment load data. In *Erosion and Sediment Transport Measurement*. IAHS Publication No 133, IAHS Press: Wallingford; 177–194.

Walling DE, Webb BW. 1988. The reliability of rating curve estimates of suspended sediment yield: some further comments. In *Sediment Budgets*, Bordas MP, Walling DE (eds). IAHS Publication No 174, IAHS Press: Wallingford; 337–350.